| Help | Home | Boolean | Advanced | Number | Order Copy | PTDLs |

[USPTO GNHDR]

(1 of 1)

United States Patent

Turtle

5,265,065

Nov. 23, 1993

Method and apparatus for information retrieval from a database by replacing domain specific stemmed phases in a *natural language* to create a search query

Inventors: **Turtle; Howard R.** (Woodbury, MN).

Assignee: **West Publishing Company** (Eagan, MN).

Appl. No.: **773,101**

Filed: **Oct. 8, 1991**

Intl. Cl. :

G06F 15/40, G06F 7/10

Current U.S. Cl.:

395/600; 364/222.81; 364/225; 364/225.3; 364/225.4; 364/226.4; 364/226.6; 364/260.4; 364/260.9; 364/282.1; 364/283.3; 364/DIG 1

Field of Search:

364/200, 300, 419, 513; 395/600

References Cited | [Referenced By]

U.S. Patent Documents

4,241,402	Dec., 1980	Mayper, Jr. et al.	364/200
4,270,182	May, 1981	Asija	364/900
4,358,824	Nov., 1982	Glickman et al.	364/200
4,384,329	May, 1983	Rosenbaum et al.	364/419
4,471,459	Sept., 1984	Dickinson et al.	364/900
4,499,553	Feb., 1985	Dickinson et al.	364/900
4,554,631	Nov., 1985	Reddington	364/300
4,580,218	Apr., 1986	Raye	364/300
4,650,848	Jun., 1987	Schramm	364/513
4,688,195	Aug., 1987	Thompson et al.	364/300
4,706,212	Nov., 1987	Toma	364/900
4,787,035	Nov., 1988	Bourne	364/300
4,823,306	Apr., 1989	Barbic et al.	364/900
4,839,853	Jun., 1989	Deerwester et al.	364/900
4,862,408	Aug., 1989	Zamora	364/900
4,868,750	Sept., 1989	Kucera et al.	364/419
4,914,590	Apr., 1990	Loatman et al.	364/419
4,918,588	Apr., 1990	Barrett et al.	364/200
4,931,935	Jun., 1990	Ohira et al.	364/419
4,972,349	Nov., 1990	Kleinberger	364/900

4,974,191	Nov., 1990	Amirghodsi et al.	364/900
4,991,087	Feb., 1991	Burkowski et al.	364/900
5,099,425	Mar., 1992	Kanno: Yuji et al.	364/419
5,109,509	Apr., 1992	Katayama et al.	395/600
5,117,349	May, 1992	Tirfling et al.	395/600
5,123,103	Jun., 1992	Ohtaki et al.	395/600

Other References

Turtle et al., "Evaluation of an Interence Network-Based Retrieval Model", Transactions on Information Systems, Association for Computer Machinery, vol. 9, No. 3, pp. 187-223 (Jul. 1991).

Croft et al., "Interactive Retrieval of Complex Documents", Information Processing and Management, vol. 26, No. 5, pp. 593-613 (1990).

Haynes, "Designing a System for the Specialized User: A Case Study", Proceedings--1985 National Online MeetingLearned Information Inc., pp. 205-213, (Apr. 30, 1985).

Croft et al, "A Retrieval Model Incorporating Hypertext Links", Hypertex '89 Proceedings, Association for Computer Machinery, pp. 213-224 (Nov. 1989).

Turtle et al, "Inference Networks for Document Retrieval", Coins Technical Report 90-07, University of Massachusetts (Mar. 1990).

Turtle et al, "Inference Network for Document Retrieval", Sigir 90, Association for Computing Machinery, pp. 1-24 (Sep. 1990).

Turtle, "Inference Network for Document Retrieval", Ph.D. Dissertation, Coins Technical Report 90-92, University of Massachusetts (Oct. 1990).

Turtle et al, "Efficient Probabilistic Inference for Text Retrieval", Riao '91 Conference Proceedings, Recherche d'Information Assistee par Ordinateur, Universitat Automa de Barcelona, Spain, pp. 644-661 (Apr. 1991).

Porter, "An Algorithm for Suffix Skipping", Program, vol. 14, pp. 130-137 (1980).

Abstract

A computer implemented process for creating a search query for an information retrieval system in which a database is provided containing a plurality of stopwords and phrases. A *natural language* input query defines the composition of the test of documents to be identified. Each word of the *natural language* input query is compared to the database in order to remove stopwords from the query. The remaining words of the input query are stemmed to their basic roots, and the sequence of stemmed words in the list is compared to phrases in the database to identify phrases in the search query. The phrases are substituted for the sequence of stemmed words from the list so that the remaining elements, namely the substituted phrases and unsubstituted stemmed words, form the search query. The completed search query elements are query nodes of a query network used to match representation nodes of a document network of an inference network. The database includes as options a topic and key database for finding numerical keys, and a synonym database for finding synonyms, both of which are employed in the query as query nodes.

46 Claims, 13 Drawing Figures

[USPTO] [CNIDR]